Summary of Proposed Metadata Standards for Utah Government

January 18, 2001

The State Library Division has prepared these standards with assistance from State Archives for the Information Technology Policy and Strategy Committee.

I Definition

"Metadata" is information about other information. A good example of metadata is a library catalog card for a book about dogs. It provides information about the format, contents and location of the book. When applied to web pages, the term describes the content, format, and contents of the information contained on the web page. "Metatags" are individual fields of metadata in document headers that provide specific information about the web document. Metadata is part of html code and is not visible to viewers, although it is used by many commercial search engines and by more specialized engines, such as Yeehaw.

II Rationale

The move from print publishing to web access for government publications creates a paradoxical challenge: while publications are more accessible than ever, they are lost in an ever-deepening sea of information. The proposed standard will allow

Increased relevancy and ranking of search results. Consistent application of metadata to web pages coupled with an appropriate search engine provides more precise, relevant, descriptive, and meaningful results to searchers. Many search engines allow each metadata field to be searched separately or in combination to achieve maximum precision.

Creation of a Yahoo-like topic directory based on the content of the metatags. With many search engines, web pages can be classified into a browsable directory, thus increasing access, as many prefer the directory approach to the raw search.

The scheduling of web publications for archival preservation using "date stamp" metatags.

Metatagging allows access to non-html and non-web based material. Metatags can be applied to "pointer records" which can identify things like pictures, sound files, and print publications for access by search engines.

III Current Architecture and Standards

Utah State web standards have consistently promoted the use of metatags, as well as comprehensible "Title" tags. The list of recommended tags below includes those mentioned in current state standards as well as several others which serve diverse but important purposes. Standardization and maximal use of metatags is important if their value is to be fully appreciated.

IV The Recommended Tags and Metatags

Tag
Title
All pages should have a "Title" tag that clearly expresses the content of the page.

Metatags
author
The full name of the agency / division / office responsible for either the creation of the document or the information contained within the document. The use of this tag has two main benefits: 1) search results can easily and clearly include the responsible agency, and 2) web documents can be searched and limited by agency, division, or office, regardless of the number or name of domains utilized by each agency.

description
A brief narrative sentence that summarizes the content and purpose of the information on the web page. This can appear in the search results, making them more comprehensible.

keywords
These are words or phrases that refer to the contents of the page in question. Keywords can be selected on several criteria including 1) words or phrases which do not appear on the page, but nevertheless reflect its content, 2) words or phrases which appear on the page and are expressive of its content. Keywords may be duplicated for emphasis, although some search engines penalize the website for doing this by placing it lower in a results list.

subjects
Several states have developed and share a list of subject headings for use in this tag. When used properly these terms assist in automatically classifying the web page into a Yahoo-like directory for additional access by the user.

pubDate
This is the date the publication is placed on the web.

docType
This tag describes what type of information the web page provides, such as directories, instructions, laws and regulations, etc.

govType
This tag describes the level of governmental entity of the producer of the information on the page. This allows for distinction between state and local governments, as well as non- and quasi-governmental agencies.

dispositionDate
This is the year the publication is expected to be removed from the web. This will be used to help preserve those non-print publications that disappear from public access when they are removed from the web and there is no analogous print version deposited in State Library Depositories.

medium
This tag describes what type of page is being accessed (html, Word file, Excel spreadsheet, picture, animation, etc.)

V Training, Metatag revision, etc.

Training and Assistance
The State Library Division can provide individualized or group training in the utilization of metatags. In addition, we can assist in the selection of metatags, and can actually prepare the metatag set for use on given pages.

Metatag Revision
From time to time it may be necessary to revise the metatag list. This revision could include 1) changing the name of the metatag, 2) eliminating or adding additional metatags, or 3) changing

the scope of a given metatag. The State Library Division will establish a procedure for accomplishing this, which will include collecting input from information users and content producers, and working the CIO to disseminate the new revisions. We would hope to use available technology to automate the change as much as possible.

Further Resources

Demonstration search engines:
      Site Server: http://168.179.144.187/WebApp/search.asp
      Compass: http://198.239.85.136:80/compass

State GILS Information Website : http://www.utah.org/gils/

Craig Neilson or Ray Matthews, at the State Library Division